DOCUMENT RESUME

ED 319 755 TM 014 854

AUTHOR Bonk, Curtis J.; Reynolds, Thomas H.
TITLE The Development of Children's Writing Awareness and Performance within a Generative/Evaluative Computerized Prompting Framework.
SPONS AGENCY American Psychological Association, Washington, D.C.; Wisconsin Alumni Research Foundation, Madison.
PUB DATE Apr 90
NOTE 69p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Child Development; Cognitive Development; Comparative Analysis; *Computer Assisted Instruction; *Cues; *Elementary School Students; Evaluation Methods; Grade 6; Grade 7; Grade 8; High Achievement; Instructional Effectiveness; Intermediate Grades; Junior High Schools; Low Achievement; Writing Improvement; *Writing Instruction; Writing Skills

ABSTRACT

The use of a generative and evaluative computerized prompting framework to improve the teaching of writing skills was studied with 164 children in grades 6, 7, and 8 with high or low writing skills, who were randomly assigned to control or treatment groups. The treatment group received computer prompts, think sheet scaffolds, and expert modeling of writing prompts. Dimensional and holistic assessment of written products, internalization of writing control strategies as measured by the Index of Writing Awareness (IWA) (an instrument developed to measure cognition about writing), open-ended questions, and a prompt sort task failed to show any advantage from treatment. Students' holistic scores were, however, significantly correlated with their awareness of writing skills. Keystroke analysis indicated that the availability of prompts seemed to lower the number of changes students made during the writing process. Differences in holistic performance, IWA performance, metacognitive guidance in writing, surface revisions, and repositioning within papers between grades 6 and 7 were not found between grades 7 and 8, suggesting that a developmental change had occurred between grades 6 and 7, and that procedural assistance may be most appropriate in grade 6. Nine tables present data from the study. A 52-item list of references and 7 appendices, including a proposed model of composition and the study instruments, are included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

CURTIS J. BONK

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED319755

THE DEVELOPMENT OF CHILDREN'S WRITING AWARENESS AND PERFORMANCE WITHIN A GENERATIVE/EVALUATIVE COMPUTERIZED PROMPTING FRAMEWORK

by

Curtis J. Bonk
West Virginia University
Dept. of Educational Psychology
609 Allen Hall
Morgantown, WV 26505

and

Thomas H. Reynolds
Educational Sciences Building
University of Wisconsin-Madison
1025 W. Johnson
Madison, WI 53706

The research reported here was supported in part by dissertation research grants obtained by the first author from the American Psychological Association (APA) and the Wisconsin Alumni Research Foundation (WARF). The authors wish to thank Gary Davis, Steve Yussen, Richard Lehrer, Joan Littlefield, Mary Gomez, Jim Middleton, Todd Franke, Robert Gilpatrick, Nancy Holty, Cheryl Hofstad, Cathy Salapa, and Barb Seeboth for their support and guidance with this project, and the many other students and teachers who made this study possible.

BEST COPY AVAILABLE

Paper presented at the Annual Meeting of the American Educational Research Association, April, 1990, Boston, MA.

TM014854

Abstract

A total of 164 children in the sixth-, seventh-, and eighth-grades were blocked according to low and high writing ability, and then randomly assigned to one of two groups: (1) a control group or (2) an experimental group which received computer prompts, think sheet scaffolds, and expert modeling of prompts. Results from dimensional and holistic assessments of written products indicated that students did not improve their writing performance as a result of exposure to prompts over a six week period. Internalization of writing control strategies as measured by the Index of Writing Awareness (IWA), open-ended guidance questions in writing, and a prompt sort task all failed to favor the treatment group. As expected, though, students' holistic scores in writing were significantly correlated with their awareness of writing skills and strategies. In addition, keystroke analyses revealed that students in the treatment group lowered the number of surface level changes they were making per 100 words of text when the prompts were available, while, at the same time, increasing their textual scanning/repositioning within text. No differences, however, were found in meaning changes as a result of these altered revisionary practices. Surface changes were negatively correlated with students' holistic scores, while repositioning within text was positively correlated with holistic scores. A computer attitudes questionnaire indicated that the treatment group considered the computer more helpful than did the control group in creating a sense of audience and in providing opportunities to evaluate their own compositions.

Grade and ability differences were also compared. High-ability students performed significantly better than low-ability students on all the performance measures, metacognitive indices, and revisionary analyses. Grade differences between sixth- and seventh-graders in holistic performance, the IWA, metacognitive guidance in writing, surface revisions, and repositioning within papers were not replicated in seventh- to eighth-grade comparisons. Thus, it appeared that a developmental increase in metacognitive writing skills and strategies had occurred between the sixth- and seventh-grades resulting in improved writing performance and revisionary practices. Grade differences on various measures provided evidence that procedural assistance may have been most appropriate at the sixth-grade level.

The Development of Children's Writing Awareness and Performance Within a Generative/Evaluative Computerized Prompting Framework

Introduction

Recent national concern over the decline in students' writing abilities has compelled educators to seek fresh approaches to teaching writing skills. Educational theorists have moved beyond the perspective of writing as a simple, technical skill to recognize the possible linkage between writing abilities and thinking abilities (Applebee, 1982; 1984; Bereiter and Scardamalia, 1985; Flower, 1987). More specifically, they propose that just as higher-order thinking involves the integration of **convergent and divergent** thinking strategies, so, too, effective writing is aided by proficient use of **generative and evaluative** composition strategies (Flower & Hayes, 1981; Scardamalia & Bereiter, 1985). This theoretical insight could significantly alter the way writing skills are taught. To date, however, there have been few assessments of the models and theories that detail writing as a generative\evaluative compositional process, or of the effectiveness of teaching strategies that stem from this theoretical perspective.

To initiate evaluation of these models, some researchers have investigated the effectiveness of prompts on children's writing (Bereiter & Scardamalia, 1982; Daiute, 1985). In these studies, prompting strategies, termed "procedural facilitators," allowed the learner to approach the writing task by using strategies that involved higher-level aspects of their problem solving processes. But, as yet, a theory of prompting for revision and/or generation of text does not exist. The goal of this research project was to

build on prior work while attempting to develop and test a theory of prompted writing. The procedural facilitators to be used in the proposed study are grouped by thinking skill; generative and evaluative prompts comprise the major classifications. Evaluative prompts, which address the logical flow, relevancy, and consistency of ideas within and between statements, aim to improve sentence clarity and the overall organization of the written text. Generative prompts, suggesting divergent new ideas and unique perspectives, encourage the learner to creatively change directions and extend ideas and concepts.

Researchers have used various methods for tracking the developmental change of students both within and between writing assignments (Bridwell, Nancarrow, & Ross, 1984). While Salomon (1988) and others (Woodruff, Bereiter, & Scardamalia, 1981) have investigated random and mandatory prompting treatments, respectively, the system proposed in this study allows the writer to choose a prompt during naturally occurring reflective moments in the composing process. The system encourages learner control in order to motivate students' writing processes and foster the internalization of executive plans and goals within their writing. In addition, the system tracks student writing progress and revisionary practices through a keystroke mapping program.

Theoretical Framework

The design of this study is best described as a problem solving approach to writing. The following researchers also have discussed this perspective with parallel terms and metaphors:

> The kind of theory we envision views writing as a process of generating and editing text within a variety of constraints (Collins & Gentner, 1980, p. 52).

> Writers and teachers have long argued that one learns through the act of writing itself, but it has been difficult to support this claim in other ways. However, if one studies the process by which a writer uses a goal to generate ideas, then consolidates those ideas and uses them to revise or regenerate new, more complex goals, one can see the learning process in action (Flower & Hayes, 1981, p. 386).

The above quotes are a reflection of the shift in the field of writing that occurred during the 1970's (Barritt & Kroll, 1978). At that time, predominate product views of writing gave way to cognitive process models. The product-oriented linear stage view of composing was replaced by a recursive model that identified planning, text generation, and revision as processes that took place at any moment in the writing process. One of the most inclusive and referenced cognitive perspectives on writing is the Flower and Hayes (1981) cognitive model of the composing process. The key assumption behind the Flower and Hayes model is that a writer does a great deal of thinking and problem solving at various levels of abstraction when composing. As the writing task or situation becomes more difficult, the writer needs additional problem solving skills and strategies to manage it (Hayes & Flower, 1986). The planning, translating, and reviewing processes, and the organizing, goal setting, generating, evaluating, and editing subprocesses of their model are not distinct stages to be completed, but are descriptions of the thinking processes that occur at any moment in the composing process (Flower & Hayes, 1981). This is a critical idea since it furnishes a framework for cognitive psychologists to speculate about

writing and reasoning and also an extensive research agenda. According to the Flower and Hayes model, the most obvious act of planning is the "generation" of ideas (Flower & Hayes, 1981). The writer's plans can be in the form of pointers, markers, words, images, or goals (Hayes & Flower, 1986). The time spent planning for text generation is extremely critical to the quality of the final text. In fact, the research agenda of Flower and Hayes and their colleagues recently concentrated on the planning procedures of experienced and inexperienced writers (Carey, 1988; Haas, 1988). Collins and Gentner (1980) detailed ways to capture ideas, including making analogies, forming critical case comparisons, defining examples, elaborating on dimensions, addressing opposite situations or dissimilar cases, and simulating a situation through overt actions or speculations. Yet, at some point during planning and idea generation, one's ideas must be conceptually organized and revised (e.g., lists, categories, divisions, or concepts maps), since conceptually-oriented text generation enhances the caliber of plans and eventual quality of text (Bereiter & Scardamalia, 1982).

More demanding than generating propositions for young writers is to verify that the text they have generated is translated into a coherent and meaningful whole. The coherency check on ideas requires text evaluation, revision, and the generation of new plans. Efficient evaluation strategies operate not only on the actual text produced, but also on internal plans or pretext (Flower et al., 1986). Thus, analysis of revisionary strategies is difficult, because experienced writers revise and reorganize not only surface

or written text, but also ideas formed or in the process of forming in their heads--pretext. The efficient writer's task representation is not static at any point in composing, but changes through extended reinterpretations of plans and goals (Flower, 1987). In fact, Bereiter and Scardamalia advocated the use of the term **"reprocessing"** in place of revision (Scardamalia & Bereiter, 1986). Reprocessing was seen as more appropriate since it accounted for whatever was produced in an episode (e.g., text, notes, or new thoughts) which then can be used in future processing cycles. As a result, reprocessing does not add something new to the text, but "transforms it."

Flower and Hayes (1981) also agreed that each word added to text or pretext constrains or limits future alternatives. They also noted that writers consistently have a "re-seeing" or "re-visioning" of the text while composing. Writing is recursive, since both planning and revising go on throughout composing. As writers reflect on what they have written and attempt to rework their thoughts and ideas, they generate and regenerate goals and ideas. Hence, the discovery of a new idea can initiate revision somewhere in the actual text or in the pretext, and this revision can lead to major and/or minor new ideas. But, perhaps, even during idea production, evaluative processes are needed to decide if the new thought, concept, plan, or image is appropriate for the task.

It is conceivable that the generative and evaluative subprocesses of writing are interdependent, even though they can be individually applied (Collins & Gentner, 1980). Collins and Gentner

also suggested that the systematic switching between these two processes is what separates good from poor writers. In addition, Flower and Hayes (1981) underscored the fact that the subprocesses of generating and evaluating are somewhat unique--only these two processes can interrupt any other process at any time during composing. Or in their words:

> First, as our model of the writing process describes, the processes of **generate** and **evaluate** appear to have the power to interrupt the writer's process at any point--and frequently they do. This means that new knowledge and/or some feature of the current text can interrupt the process at any time through the processes of **generate** and **evaluate**. This allows for a flexible collaboration among goals, knowledge, and text. Yet this collaboration often culminates in a revision of previous goals (Flower & Hayes, p. 380).

This notion of an interaction between generative and evaluative processes in planning and revising is elaborated both in the following paragraphs and also in the discussion of the experimenter's proposed model of the cognitive processes in writing presented later in this paper (**see Appendix A**).

Collins and Gentner (1980) noted that elementary students lack these important writing skills when they first attempt to write expository text because they are unable to concentrate or voluntarily interrupt their writing processes to focus on critical aspects of text. The findings from Collins and Gentner are interesting since they draw attention to the executive or monitoring component of the Hayes and Flower model. According to the Flower and Hayes (1981) model, the successful writer is one who is able to monitor and direct his/her own success. But what does this mean?

As indicated, young children's difficulties in generating and evaluating text lies in their lack of an executive routine to handle the switching between writing processes (Flower & Hayes, 1981). A high level executive routine is the most valuable component of text production, since it directs and regulates the whole writing process (Bereiter, 1980). Besides the regulation of processes, children need knowledge about the capabilities of their own thinking and regulatory processes (Flower, 1987). In effect, the monitor is useful both for the encouragement of new ideas and also for the evaluation of them. Even when children have concept generation and rewriting skills, the excessive demands of the writing situation inhibits their use (Flower et al., 1986). Monitoring their changing plans, goals, and other constraints within the limits of current working memory is even more difficult (Flower, 1987). Therefore, those intervening into inexperienced writer's executive scheme or strategies must consider these limitations and adapt their treatments to them.

Bereiter and Scardamalia (1985) suggested that elementary children have cognitive coping strategies that get them through tasks without the need for effortful higher-order thinking or executive control processes. Children often do not have the benefit of feedback from a conversational partner when they enter the domain of writing (Bereiter, 1980). To illustrate, Bereiter and Scardamalia (1982) referred to the implicit and explicit cues from a listener within a conversation that indicate to the speaker that she should proceed, stop, elaborate, define, list, generate

alternatives for, or refine ideas. In stark contrast, the writing feedback loop must function autonomously. The young writer must find a way to adapt to radical changes affecting her involvement with both her own thought and her immediate audience.

Without support, the text produced from young writers appears egocentric. According to Bereiter and Scardamalia (1985; 1987), young children's writing is immature because they are operating within the **"knowledge telling"** problem space. Here, children concentrate mainly on the words they want to say without framing them according to discourse conventions or audience expectations (Scardamalia, Bereiter, & Steinbach, 1984). One signal of text written in the knowledge telling mode is content written in the order generated. Also, knowledge tellers will reexamine assignment cues for topical prompts when stuck, rather than inspect internal plans and goals (Scardamalia & Bereiter, 1982). The slightest prompting persuades knowledge tellers to write more since their major goal is to get content on paper or externalize their existing knowledge (Bereiter & Scardamalia, 1985).

Bereiter and Scardamalia (1982) posited an intermediate stage in writing development between knowledge telling and more mature **"knowledge transforming."** In this stage, the young writer depends on concentrated self-cuing or outside intervention in order to generate extended written discourse while using conversation-related schema structures. But, initially, the mechanical constraints of text production drain available cognitive resources, eliminating opportunities for self-questioning or internal dialogues to develop

(Bereiter & Scardamalia, 1987). Therefore, these researchers argued that for young children to consider rhetorical constraints, such as audience, in a similar fashion to the cues and constraints provided by a conversational partner, they need support strategies.

Recently, McCutchen (1988) made a critical distinction between automaticity and functional automaticity/fluency in writing. She concluded that the major difference between knowledge tellers and skilled knowledge transformers was their metacognitive control over their writing fluency subprocesses. In addition, she argues that highly automatized procedures can be costly since they are resistent to intervention and are difficult for the writer, teacher, or researcher to analyze. McCutchen proposed that writing skill interventions must attempt to strike a balance between fostering automaticity/fluency and encouraging metacognitive awareness and control of the writing process. For example, she suggested that a mild form of intervention such as procedural facilitation may be one way to overcome "content spills," while fostering metacognition in writing.

Important to this study, McCutchen argued that using procedural facilitation where rule systems are too complex to teach (e.g., writing) would promote self-regulatory mechanisms (Bereiter and Scardamalia's, 1987). This type of intervention technique oversees the overall executive procedure, while providing cues or routines for switching the writer between text generation and revision. Scardamalia and Bereiter (1985, p. 566) defined **procedural facilitation** as "routines and external aids designed to reduce the

processing burden involved in bringing additional regulatory mechanisms into use." Basically, procedural facilitation was designed as a tool for developing children's metacognition in writing (Bereiter & Scardamalia, 1987). Procedural facilitation provides young writers with temporary supports, similar to those they previously received in conversation, thereby encouraging them to adopt more complex writing strategies.

One of the key goals of the prompts is to help the child become more reflective when approaching a writing task. Procedural assistance can offer encouragement in content generation, keeping the student on topic, generating appropriate plans and goals, attending to structural elements in the text, diagnosing and operating on textual inconsistencies and problems, producing a coherent whole, and, more generally, increasing the level of sophistication in writing (Scardamalia & Bereiter, 1985). In effect, covert mental operations are made overt. Procedural facilitation is contrasted with substantive facilitation in which the outside support actually reduces part of the task burden or intellectual demands of the task. In substantive facilitation, comments, cues, probes, or even hints are directed specifically at something the student has written (i.e., they are topic-specific), whereas in procedural facilitation the cues are extremely general and nonspecific, but, nonetheless, are familiar to the writer.

The prompts used in procedural facilitation techniques act as a partner would in a conversation, enabling the writer to decenter from personal thoughts and think about potential readers. This

awareness of the reader or social-cognitive ability may be one of the most difficult aspects of effective writing (Bonk, 1990; Martlew, 1983). By forcing the writer to reflect on his/her writing routines, the prompts may encourage automated processes to come under review and eventual control (McCutchen, 1988). What the technique essentially facilitates is an internal dialogue, or at least some form of assisted monologue (Scardamalia et al., 1984). The interaction between one's actual text and intended text when highlighted through procedural facilitation is an invitation for reflection. Pausing to internally reflect and discuss textual inconsistencies or audience needs is a valuable writing skill (Martlew, 1983). However, the internal dialogue of the expert is hard to find in the verbal protocols of novices (Daiute, 1985; Scardamalia et al., 1984). Therefore, the intent of signals sent by external sources such as procedural prompts is to encourage the child to work within an externally supported, interactive writing environment that eventually functions as an independent inner dialogue.

To initiate evaluation of interactive models of the composing process, some researchers have investigated the effectiveness of prompts on children's writing (Bereiter & Scardamalia, 1982; Daiute, 1985). They have given students sentence openers (e.g., "I think...," or "One reason is...,"), contentless prompts (e.g., "Can you write some more?"), and planning and goal-setting prompts (e.g., "give an opinion") (Scardamalia & Bereiter, 1986; Woodruff, Scardamalia, & Bereiter, 1981). But, as yet, a complete theory of

prompting for generation, revision, or reprocessing of text does not exist. In addition, though researchers allude to the importance of switching between generative and evaluative aspects of text, procedural facilitation research has yet to incorporate both generative and evaluative processing support. Research from the Flower and Hayes group has addressed planning (Carey, 1988; Haas 1988), while others have focused mainly on evaluative and revisionary procedures in writing (Bereiter & Scardamalia, 1987; Daiute 1985; 1986; Scardamalia et al., 1984; Woodruff et al., 1981). For a review of preliminary findings regarding procedural facilitation systems (prompted writing) see Bonk (1989).

Bereiter and Scardamalia (1987) argued that procedural facilitation was one method to lead students on a path of active, self-regulated learning. When students are stuck or inactive, they can use the procedural prompts to get more involved in the activity. Thus, there must be conscious and purposeful use of this strategy. Additionally, the more students understand of the strategy (e.g., through expert modeling and peer support), the more likely they will feel a sense of ownership in it and be motivated to use it. In order to maximize the instructional effectiveness of the intervention, procedural facilitation is aimed at the upper limits of the child's learning envelope or growing edge of competence.

Some researchers on children's writing development have begun to think seriously about the implications of Vygotsky's notion that all good learning may occur in advance of development (Vygotsky, 1986). Writing instruction that aims for the upper limit of the

child's zone of proximal development where young writers witness expert questioning, includes both teacher conferences, expert modeling, peer response groups, and prompt internalization relate recent writing research to Vygotskian psychology (Daiute, 1985; DiPardo & Freedman, 1988; Englert & Raphael, 1988; Salomon, 1988). Freedman (1987) pointed out that when applying Vygotskian theory to writing, the intervention should include: (1) collaboration between the writer and someone with more writing expertise; (2) assistance in solving writing tasks and problems not possible by oneself; and (3) scaffolds which self-destruct when they lead to enhanced independent problem solving in writing.

Salomon (1988) also used Vygotsky's ideas of zones of proximal development, mediated instruction, and internalization to conceptualize tool-based instruction. He claimed that interactions with a supportive partner could include a computer, because this tool is so important to interpersonal processing (i.e., they change one's relationship to the task and to the world). The intelligence displayed (or seemingly displayed) by computer tools can be internalized by the user, leaving some sort of "cognitive residue." Yet, internalization is not recopying of the process displayed, but involves reconstruction of the activity according to one's perceived needs. For a tool to be a prime candidate for internalization, it should: (1) present opportunities not commonly encountered elsewhere; (2) be explicit in its "intelligent" operations; and (3) use strategies that are generalizable and capable of being carried out in one's mind, while the user must have the ability to use

metacognitive skills to govern independent use of the strategy.

A Model of Generating and Evaluating Text

According to Caccamise (1987), the particular writing process that should be the main concern of researchers is planning. As mentioned earlier, this is supported by the fact that 80% of composing time is typically spent planning. Caccamise's own investigation into planning and idea generation has demonstrated that when a student receives or creates a topic, a retrieval probe is generated to search preexisting world knowledge of the topic. At the same time, an immediate editing process develops according to task familiarity. The more familiar the topic, the greater the number of ideas generated. The size of the search set also depends on the task specificity; the more general the probe, the greater the number of possible concepts and ideas generated (though they may lack tight organization).

A similar pattern holds for revision as for idea generation. Revision is a central part of writing (Fitzgerald, 1987) which extends considerably beyond mechanical editing (Bartlett, 1982). Revision must be viewed as a separate process from translation, since it involves comparing actual text with intended text and operating on those differences. Because revision often leads to new thoughts and ideas, it is seen as vital to the generative plans and goals of writing. Thus, revision involves both generative and evaluative processes. Just as Caccamise (1987) pointed out that there are numerous constraints during planning and idea generation, similar constraints and situational variables affect revision and

resulting writing quality (e.g., genre, task familiarity, subject, and length) (Witte & Faigley, 1981).

There are many differences already noted in the revising strategies between experts and novices. Experts make more revisionary changes that affect the meaning of the text, while inexperienced writers focus on the local level by making spelling, word, or other surface level changes (Witte & Faigley, 1981). Further, while experts wait to correct most mechanical errors until after the topical criteria have been satisfied, inexperienced writers stop writing and reevaluating the gist of their text in order to correct minor errors, but not meaning. As with planning, novices must be made aware of larger textual and rhetorical revisionary concerns.

As noted in the author's model displayed in Appendix A, both generative and evaluative processes are critical to planning and revising text and pretext, though in different strengths. Note that the emphasis remains on idea generation in the planning phase and on the evaluation of those ideas in the revision phase. However, metacognitive control is needed in order for the evaluative processes to have a significant role in text generation and for revising to be seen as a generative process. Uncontrolled processes can have uncertain consequences in writing (McCutchen, 1988). In Appendix A, writers operating at the lower (spontaneous) level are generating text according to content goals or length requirements, without an interaction occurring between generative and evaluative processes in either their plans or revisionary tactics.

Procedural prompts perform a critical function in refocusing students' attention to the interaction between generative and evaluative processes. Students should be most concerned with the coherence, consistency, elaborateness, and logical sequence of their ideas, not the mechanics (e.g., spelling) (Bereiter & Scardamalia, 1987). They must be made aware of the interconnectedness of their thoughts; each change made within the created text or pretext impacts on the plans for every other thought, image, or note. As mentioned, their ideal textual environment is one of reprocessing (Fitzgerald, 1987; Scardamalia & Bereiter, 1986), not spontaneity.

When elementary students are permitted to operate in the spontaneous level of knowledge telling, they "downslide" or get pulled into the local demands of topic-directed content generation and/or editing for mechanical mistakes (Bruce et al., 1982). Interventional tactics using computers might refocus children's attention to a more "principled" or metacognitive level (Comeaux & Lehrer, 1986). In the model presented in Appendix A, generative and evaluative prompts are noted as one way to shift children's attention to higher levels of processing and control in the writing hierarchy. In this model, metacognitively aware writers organize and reorganize textual and pretextual plans and goals. Thus, internalization of the procedural strategies enables the knowledge teller to move toward a more mature stage of writing.

Editing and content generation are not entirely neglected processes; they are still available when needed. For example, the writer may, at times, have thoughts flowing so fluently that she

would want to freely write them down without the additional demands of evaluation. At some point, however, she would come back to analyze these thoughts for overall coherence. Thus, the mature writer experiencing spontaneous text generation or stream of consciousness maintains some degree of deliberateness or control over actual text produced.

The proposed model indicates that there are hierarchical levels within writing that function together to enhance writing quality. The **generative and evaluative** organizing and reorganizing of one's plans and goals is the critical difference noted in this model of writing. Without deliberate attempts to generate and evaluate plans and text, writing would not exist. Yet, as evident throughout the literature review, this interconnection between generative and evaluative processes is often missing in children's writing.

The present study attempted to repair the missing processes by exploring the effectiveness of computerized generative and evaluative prompts within the development of children's writing. **Development was explored in six different ways.** First, midtreatment and post-treatment writing quality data were compared between the treatment group (prompted condition) and the control group. Second, development was analyzed between three grade levels, sixth-, seventh-, and eighth-grades. Third development was compared between high and low ability students. Fourth, development within an individual writing assignment was tracked and reevaluated through a keystroke mapping system. Fifth, possible internalization of prompts was examined. Finally, children's writing gain scores over

a ten week period were compared.

Methods/Techniques

A total of 164 middle school students, 53 sixth-, 52 seventh-, and 59 eighth-grade students [Merton Unified School District #9][1] took part in a 13 week course (one session/week) on electronic writing using a word processing package entitled "WordPerfect 4.2," from WordPerfect Corporation. In the initial three weeks (or three sessions), students were given preliminary training on WordPerfect during November, 1988. After preliminary training, they were blocked according to low and high writing ability, and then randomly assigned within each grade to one of two groups: (1) a control group, or (2) an experimental group which received computer prompts, think sheet scaffolds, and expert modeling of prompts. The experimental group viewed a twenty minute videotape of the two expert college writers, one male and one female, modeling how to use the prompts. Next, there was a live five minute modeling session of the computerized prompting program. After the modeling, students were given a chance to practice the alt-shift letter combination needed to invoke a prompt. Finally, the experimental group received a list of the generative and evaluative prompts for referential and recording scaffolding purposes as an additional mediational feature. Students in the experimental groups were asked to use the prompts at least 12 and 10 times, respectively, during the two sessions of paper two; 8 and

[1]These students represented the entire enrollment of Merton at these grade levels.

6 times, respectively, during the two sessions of paper three; and when needed during the two sessions of paper four--the final exposure to the prompts. More specifically, they were asked to use them during reflective moments in the composing process. Of the ten weeks of word processing in the Spring, the prompts were available only during weeks 3-8 (see Table 1 for a visual of the procedure).

A blocking variable on writing ability was incorporated into the design based on the average of each student's reading and language scores from the Metropolitan Achievement Tests. The blocking variable was added to determine the effects of prompts on low- and high-ability students.

The procedural facilitators used in the study were grouped by thinking skill based on the cognitive process model of Flower and Hayes and the author's model of writing explained previously; generative and evaluative prompts comprised the major classifications (See Appendix B). Generative prompts, suggesting divergent new ideas and unique perspectives, were developed to encourage the learner to creatively change directions and extend ideas and concepts. The four generative prompt categories were derived from the literature on creativity (Davis, 1986; Torrance, 1974), and were as follows: (1) **fluency** (the ability to produce many types of ideas); (2) **flexibility** (the ability to take different approaches to a problem or think of different categories); (3) **originality** (uniqueness or nonconformity in thought); and (4) **elaboration** (the ability to add details or embellishments to an idea) (Davis, 1986). These same categories were subscores of the

widely used Torrance Tests of Creative Thinking (Torrance, 1974). Due to the abstractness of these terms, the names were scaffolded down to the appropriate level. Subjects knew the generative prompts as **"more ideas," "types of ideas," "new ideas,"** and **"extenders."**

Evaluative prompts, which addressed the logical flow, relevancy, and consistency of ideas within and between statements, aimed to improve sentence clarity and the overall organization of the written text. The evaluative prompting categories were as follows: (1) **relevancy** (needed information that enhances the overall theme); (2) **logic** (clear flow of the paper); (3) **assumptions** (recognizing bias in one's thoughts); and (4) **conclusions** (drawing appropriate summary statements). These categories were derived from overlapping categories within prior procedural facilitation revision studies by Daiute (1985) and Scardamalia et al. (1984) and various critical thinking projects involving writing (Talbot, 1986). Students understood the four evaluative categories as **"quality," clear/logical," "assuming,"** and **"conclusions."**

Pilot testing with four sixth-, seventh-, and eighth-graders and also with four college students evaluated prompt effectiveness. Considering the categories and the type of assignments they were designed for, it was not surprising that the prompts were more facilitative for expository text than for narratives. Prompts rated highly beneficial by subjects in the pilot research were selected for this study.

Prompts were organized from left (generative) to right (evaluative) in a matrix on a computer keyboard (see Appendix B).

Three keys running diagonally down the keyboard formed one category (i.e., the "q," "a," and "z" keys formed the fluency category). The prompts were in the form of statements and questions requesting that the writer ponder over possible textual changes (e.g., "Have you provided enough information to back up your claims and conclusions?," or "Try to broaden the focus of your paper by including exceptions to what you are saying"). The prompts were programmed into WordPerfect's macro function capabilities. When "alt-shift-letter" combination keys were pressed, a prompt appeared in the bottom two lines of the screen (twenty-two lines of text were saved for the paper). At that point, the student had the option of using the prompt for reflection, continuing with current thoughts, or invoking another prompt. Meaning and surface changes as a result of the prompts were recorded through a keystroke mapping system described later.

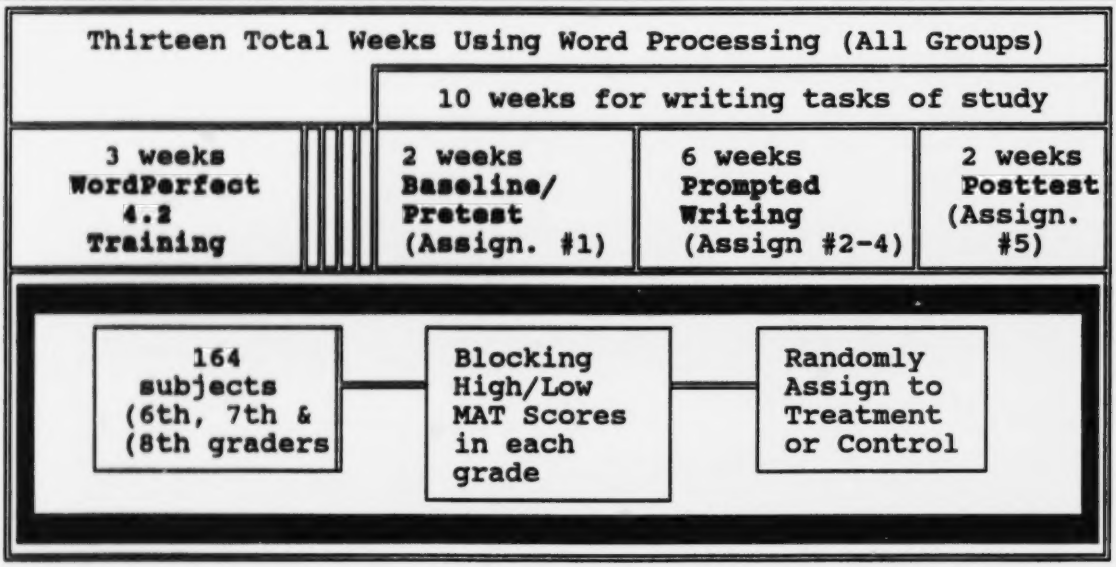
Students composed their expository texts in two sessions. Topics were assigned one week before actual composing time to allow students time to prepare. Students were told to concentrate on their content, rather than spelling or grammar. Forty-five minutes was allocated for each class, though only 40 minutes was available for composing due experimenter set up time.

After the three weeks of training, the remaining ten-week writing project consisted of three analyzable events. First of all, during the initial two weeks of training in January, writing task #1 served as baseline data. During the next six weeks (weeks 3-8), students in the experimental group had the prompts available during

composing (tasks 2-4). The final task, task #5, was assigned during weeks 9 and 10 without the availability of the prompting program. Therefore, task #5 was used as a posttest of writing ability. The pretreatment, midtreatment, and posttreatment data were used to make comparisons of the writing samples between groups, grades, and ability levels and over time (see Table 1). Expository writing assignments were chosen for all five assignments since they required generative/inventive cognitive processes (Ruth & Murphy, 1988), as well as advanced evaluative/revisionary processes (Bereiter & Scardamalia, 1987). The five writing tasks were selected according to student interest across grade levels and gender.

In similarity to Bridwell et al. (1984, 1985), a keystroke mapping procedure was used to track revision from multiple writing samples for every writer within each task. The keystroke mapping tool was designed to replay individual keystrokes at preselected speeds, thereby providing a chronological picture of student drafts and detailed documentation of within draft changes and individual progress (for more information on this procedure, see Reynolds & Bonk, 1990). Instead of analyzing these writing tasks within a "draft and revise" scheme, each paper was analyzed using the "reprocessing" framework or as one semantic unit from beginning to end. Using the replay of keystrokes and the Taxonomy of Revision Changes (Faigley & Witte, 1981), the revision factors that were analyzed in this study were: (1) **surface revisions** (a collapsed category including formal--format, spelling, grammar changes; and meaning-preserving--word changes); and (2) **meaning revisions**

Summary of Overall Methodology and Intervention Sequence


Thirteen Total Weeks Using Word Processing (All Groups)
10 weeks for writing tasks of study
3 weeks WordPerfect 4.2 Training
2 weeks Baseline/ Pretest (Assign. #1)
6 weeks Prompted Writing (Assign #2-4)
2 weeks Posttest (Assign. #5)
164 subjects (6th, 7th & (8th graders
Blocking High/Low MAT Scores in each grade
Randomly Assign to Treatment or Control

Sub-groupings (approximately 12-15 subjects/subgroup):

Treatment Low-ability (6th)*
Control Low-ability (6th)
Treatment High-ability (6th)
Control High-ability (6th)

Treatment Low-ability (7th)
Control Low-ability (7th)
Treatment High-ability (7th)
Control High-ability (7th)

Treatment Low-ability (8th)
Control Low-ability (8th)
Treatment High-ability (8th)
Control High-ability (8th)

*The treatment group members were exposed to generative and evaluative procedural facilitation prompts, expert modeling of the prompts, and thinking sheet prompt scaffolds.

Assessments were made of students writing quality (holistic and dimensional trait scoring), revisionary practices (surface and meaning changes), and awareness of writing strategies (IWA).

(text-based changes at the phrase, sentence, and multisentence level). In addition, **errant keystrokes** was added to the original Table taxonomy to separate these revisions from spelling, format, grammar, and word changes. An errant keystroke was considered any format or spelling change within a word that took place before starting another word. Prompt effectiveness also was recorded.

Since writing involves many dimensions besides revision, there was a need to combine the keystroke analysis with other research tools. Therefore, the evaluations of the written products extended beyond analyses of in-process revisions to both holistic and analytic assessments. When considering the generative/evaluative focus of the present study, Purves' Analytical Scheme for Critical Thinking Skills recommended by Hawisher and Fortune (1988) to investigate writing and reasoning skills, was ideal. The original instrument is composed of three dimensions: (1) content thinking (seven items); (2) organization (three items); and (3) style/tone (three items). To further evaluate the effectiveness of the prompts, the content thinking and organization components were reordered into generative and evaluative subscales. Item #1 (**adequacy of information**), #2 (**the richness of information**), #4 (**the number and depth of inferences made**), and #7 (**consideration of alternatives**) comprised the content thinking-generative subscale. Item #3 (**relationships drawn**), #5 (**synthesis or drawing together of ideas**), #6 (**evaluation or making judgments as to the merits of ideas**), and a combination of items 8-10 (**framing, grouping, and unity**) comprised the content thinking-**evaluative subscale**. The 465

papers in this study were rated using all eight dimensions on a scale of "1" (low on this category) to "5" (high on this category). The modification of Purves' critical thinking instrument for the purposes of this study is listed in Appendix G.

Recently, Hawisher and Fortune (1988) found that a holistic scoring framework facilitated later scoring using Purves' critical thinking scheme. A holistic rating also provided an overall impression of composition quality that the Purves instrument lacked. The current study followed Hawisher and Fortune's (1988) lead in using Purves' thinking skill measure in conjunction with a holistic instrument to judge "writing quality." The holistic instrument used in this study was developed by the Educational Testing Service (ETS). The scoring was based on a 6-point scheme (see Appendix F).

One week after the treatment period, students completed five different instruments: (1) first was a metacognitive measure which incorporated three open-ended questions regarding the giving of advice in writing (see Appendix C) (Englert, Raphael, Fear, & Anderson, 1988; Salomon, 1988); (2) one open-ended question regarding the giving of advice in reading (also in Appendix C); (3) a questionnaire regarding students' perception of the assistance provided by the computer system (see Appendix D); (4) the Index of Writing Awareness (IWA) (see Appendix E); and (5) a thirteen item prompt sort task. These instruments were administered to all students to assess the internalization of the questioning strategies and heuristics. The prompt sorting task (#5 above) required students to select questions/prompts which appeared to be similar

to each other. In this task, two points were allotted to within-category responses (e.g., both addressing fluency); one point for strictly generative or strictly evaluative responses; and no points for responses which mixed generative and evaluative prompts. The 10 item computer questionnaire (#6 above) was designed to learn whether the computer was regarded as a conversational partner or tool within any of the grades or groups. The final instrument, the Index of Writing Awareness (IWA-#4 above) (see Appendix E), was based on the Index of Reading Awareness (IRA) which was developed to measure children's metacognition about their reading skills (Jacobs & Paris, 1987). Bonk (in preparation) discusses the usefulness of the IWA in diagnosing early adolescent writing difficulties along with various instrument reliability issues.

Results

Both the metacognitive advice scores and the holistic and dimensional ratings of the papers required interrater reliability checks. All raters were trained using cases not included in the random sample. A stratified random sample of the open-ended advice questions in reading and writing were scored by another individual. Interrater correlations were based on a sample of 30 cases out of a possible of 164 cases. The scoring guidelines for these two measures are listed in Appendix C. Interrater correlations were $r = .91$ ($p < .01$) on the open-ended question in writing, and $r = .85$ ($p < .01$) on the open-ended question in reading. Based on an .80 criterion these coefficients were considered acceptable.

Ninety, or about 20% of the total 465 essays within papers 1, 3, and 5 were randomly selected and divided equally between two local writing experts. Stratified random sampling ensured an equal number of subjects selected by grade, group, and ability. None of the correlation scores met criterion. As a result, all papers were blindly reread and rescored; discrepancies from the first scoring were recorded. At that time, a third rater, who received more training than the first two raters, was hired to blindly score a new stratified sample of 90 papers (20% of the total); 30 from each assignment, once again equally split between grade, group, and ability. The correlation coefficient between the holistic scores of rater #3 and the experimenter was .83. Interrater correlations across all eight dimensions of the Purves instrument was .73. These correlations were considered acceptable.

In the primary analysis, it was expected that the treatment group would perform qualitatively better than the control group on the mid-treatment (with prompts) and post-treatment (without prompts) papers. Group by ability differences were expected to favor the treatment subgroup within each grade. This hypothesis was tested using the holistic, generative, and evaluative scores of papers three and five. Planned comparisons were performed within each grade and ability level to compare specific groups. Six specific planned comparisons were performed on each of these scores. The comparisons were designed to investigate possible treatment (experimental vs. control) differences within the low- or high-ability students at the sixth-, seventh-, and eighth-grades. The

acceptable alpha level was set at .05 for all planned and post hoc comparisons.

Comparisons were performed for both paper number three (the prompted condition), and paper number five (the posttest without prompts). Dunn directional planned comparisons were calculated on this data. Only one comparison was significant. On the posttest, the sixth-grade high-ability treatment group outperformed the corresponding control group on both the holistic scoring ($\underline{t}$ = 2.15; $p < .05$), and the evaluation dimension of the modified Purves instrument ($\underline{t}$ = 2.54; $p < .05$). The means, standard deviations, and test statistics for paper five are listed in Table 2 (see Bonk, 1989 for additional data regarding either paper one or three).

The next analysis addressed whether older students were writing holistically better papers, as predicted by the overall design of this study. To avoid redundancy, only the holistic scores were investigated for grade differences. Directional Dunn comparisons were performed on the three possible grade comparisons per paper (see Table 3). In both papers three and five, the eighth-graders outperformed the sixth-graders ($\underline{t}D$ = 6.45, and $\underline{t}D$ = 5.64, respectively). Also, in both instances the seventh-graders scored significantly higher than the sixth-graders (paper one: $\underline{t}D$ = 3.08, $p < .05$; and paper three: $\underline{t}D$ = 5.80, $p < .05$). The eighth-graders scored higher than the seventh-graders on paper three ($\underline{t}D$ = 2.87, $p < .05$), but not on paper five ($\underline{t}D$ = -.13, $p > .05$).

In order to determine whether students were adequately split the students into low- and high-ability groupings, a 2 (groups) by

Table 2

Means, Standard Deviations, and Test Statistics for the Planned Subgroup Comparisons in Paper Five (Posttest)

1. Holistic Scores:

Grade	Ability	Treatment Mean	Treatment SD	Control Mean	Control SD	Test Stat
6th	Low	1.69	(.63)	1.85	(.69)	-.34
	High	3.50	(1.56)	2.54	(1.05)	2.15*
7th	Low	2.92	(1.26)	2.77	(1.01)	.34
	High	4.31	(1.03)	4.76	(.72)	-1.01
8th	Low	3.14	(1.02)	3.18	(1.42)	-.12
	High	4.33	(1.23)	4.00	(1.62)	.77

2. Generative Scores:

Grade	Ability	Treatment Mean	Treatment SD	Control Mean	Control SD	Test Stat
6th	Low	7.23	(2.45)	7.62	(2.22)	-.24
	High	12.21	(5.58)	9.61	(3.91)	1.68
7th	Low	11.38	(5.45)	13.50	(4.38)	-.34
	High	14.84	(3.21)	16.15	(3.31)	-.83
8th	Low	10.36	(3.50)	11.19	(4.15)	-.57
	High	14.13	(3.80)	13.07	(5.09)	.71

3. Evaluative Scores:

Grade	Ability	Treatment Mean	Treatment SD	Control Mean	Control SD	Test Stat
6th	Low	5.69	(1.25)	6.62	(1.39)	-.70
	High	11.29	(5.08)	8.00	(2.45)	2.54*
7th	Low	9.31	(3.84)	8.69	(2.78)	.47
	High	14.08	(3.68)	14.31	(2.40)	-.18
8th	Low	9.86	(2.68)	10.00	(3.46)	-.12
	High	14.13	(4.24)	13.14	(4.34)	.79

1. Holistic Scores: MSW = 1.35
2. Generative Scores: MSW = 16.14
3. Evaluative Scores: MSW = 11.32

*$p < .05$; $\underline{t}D = 1.87$

Table 3

Means, Standard Deviations, and Test Statistics for the Planned Grade Comparisons on the Holistic Scores in Papers Three and Five

1. Holistic Scores: (Paper Three)

Grade	Mean	SD	Grade	Mean	SD	Test Stat
8th	3.98	(1.29)	6th	2.56	(1.21)	6.45*
8th	3.98	(1.29)	7th	3.30	(1.30)	2.87*
7th	3.30	(1.30)	6th	2.56	(1.21)	3.08*

1. Holistic Scores: (Paper Five)

Grade	Mean	SD	Grade	Mean	SD	Test Stat
8th	3.66	(1.41)	6th	2.42	(1.26)	5.64*
8th	3.66	(1.41)	7th	3.69	(1.32)	-.13
7th	3.69	(1.32)	6th	2.42	(1.26)	5.80*

1. Paper Three: MSW = 1.28; n: 6th=50; 7th=40; 8th=53.
2. Paper Five: MSW = 1.35; n: 6th=53; 7th=52; 8th=59.

*p < .05; directional tD1.5,131(.05) = 2.13.

2 (ability levels) by 3 (grades) ANOVA was performed on the holistic scores for papers three and five. Both of the main effects for ability were significant (paper three: $F1,131 = 35.87$, $p < .05$; and paper five: $F1,152 = 51.74$, $p < .05$). Thus, by comparing the means it can be concluded that high-ability students' holistic scores were significantly better than low-ability students on papers three and paper five (paper three: high $X = 3.85$, $sd = 1.38$; low $X = 2.71$, $sd = 1.17$; and paper five: high $X = 3.91$, $sd = 1.41$; low $X = 2.62$, $sd = 1.19$). Additional Scheffe post hoc analyses were performed to verify that the ability group differences also were significant within each of the three grade levels. All three ability within grade comparisons were significant in both papers, with the high-ability students at each grade level outperforming the low-ability students.

The next hypothesis addressed students' writing development over time. This hypothesis stated that over six weeks, students in the experimental group would internalize the generative and evaluative prompts. This was expected to be reflected in higher holistic and dimensional treatment change scores by each ability subgroup compared to the control subgroups. In order to determine whether any of the six treatment subgroups improved over the prompting period, pretest to posttest change ("gain") scores were obtained for the holistic, generative, and evaluative measures. Change scores were derived by subtracting the pretest score (paper one) from the posttest score (paper five) for each subject. The identical six subgroup directional Dunn planned comparisons were

performed on this data as were performed in relation to hypothesis one. None of the planned comparisons within the 12 subgroups were significant.

Dunn planned grade comparisons were also performed. The planned comparisons between the sixth- and eighth-grade holistic change scores indicated that there was a significant change in a positive direction for the sixth-graders after ten weeks of word processing compared to a negative change for subjects in the eighth-grade ($t = 2.44$, $p < .05$). Also, there was a significant difference in a positive direction for the comparison between the seventh- and eighth-graders, once again favoring the younger group ($t = 2.54$, $p < .05$). These additional comparisons were performed in order to discover at what grade levels writing practice using the word processor regardless of treatment might be most beneficial.

Additional questions regarding the internalization, transfer, and usefulness of the prompts were expected to be answered by administering the IWA, the sort task, the open-ended metacognitive guidance questions in reading and writing, and a computer attitudes questionnaire at the end of the treatment period. As with the previous two hypotheses, it was predicted that the six treatment subgroups would score higher than their matched control groups on each of these five measures. As noted earlier, each of these instruments was designed for a specific purpose. For example, the sort task was intended to find out whether there was a difference in students' internal categorization of the prompts, while the open-ended question regarding guidance in reading was included to test

for possible transfer of what was learned to reading.

The results of the pairwise comparisons for the three internalization scores--the sort task, the IWA, and the open-ended question measuring metacognition in writing are listed in Table 4. See Bonk (1989) for data regarding the planned comparisons for the other two measures: the transfer task to reading and the attitudes questionnaire. Results repeat the basic findings of the previous two hypotheses; even with directional (one tail) hypotheses, only one of the pairwise comparisons was significant. The seventh-grade high-ability treatment group scored significantly higher than its control group on the open-ended question regarding advice that they would give a friend who did not read well (the transfer task) (t = 2.11; $p < .05$; see Table 4). Though both the low- and high-ability treatment sixth-graders outscored their control counterparts on all five of these measures, the differences were not significant.

Also investigated was metacognitive development by grade and ability level. It was predicted earlier that older students would score higher than younger students on the IWA, the open-ended questions in reading and writing, and the sort task. In addition, the IWA was used to test metacognitive scores by ability level and also validate the ability blocking variable. Students of higher writing ability were expected to score significantly above lower-ability students.

The three possible directional Dunn planned comparisons (8th versus 7th; 8th versus 6th; and 7th versus 6th) were performed on

these four measures[2]. The results of these analyses are listed in Table 5. Significant effects in the hypothesized direction were found on three of the four measures. First, the pairwise comparisons between eighth- and sixth-grade and also between seventh- and sixth-grade on the IWA reached significance, with the older students scoring higher ($\underline{t}$ = 2.25, $p < .05$; and $\underline{t}$ = 2.91, $p < .05$, respectively). Yet, the comparison between eighth and seventh-graders was not significant on this measure ($\underline{t}$ = -.75, $p > .05$). The open-ended guidance question in writing gave similar results: eighth-graders scored better than sixth-graders ($\underline{t}$ = 5.84, $p < .05$); seventh-graders scored better than sixth-graders ($\underline{t}$ = 3.83, $p < .05$); and there were no significant differences between the eighth-grade and the seventh-grade subjects ($\underline{t}$ = 1.90, $p > .05$). However, this time the differences between the older groups favored the eighth-graders and approached significance ($p < .10$). The open-ended question in reading findings were also somewhat similar: eighth-graders outperformed the sixth-graders ($\underline{t}$ = 2.20, $p < .05$); differences between the seventh-graders and the sixth-graders approached significance ($\underline{t}$ = 1.87, $p < .10$); and the means of the seventh- and eighth-graders were not significantly different ($\underline{t}$ = .28, $p > .05$). Finally, though not shown in this table, there were no significant differences between grades to report regarding the sort task.[3]

---

[2]Slight differences in the number of students included in each of these grade analyses was due to a few students failing to complete all four instruments.

[3]As expected, higher-ability students significantly outperformed lower-ability students on all of these measures.

Table 4

Means, Standard Deviations, and Test Statistics for the Planned Subgroup Comparisons Within the Internalization Measures

Grade	Ability	Treatment Mean	Treatment SD	Control Mean	Control SD	Test Stat
1. The Index of Writing Awareness:						
6th	Low	25.46	(3.73)	25.00	(3.81)	.26
	High	28.43	(4.73)	26.15	(5.51)	1.33
7th	Low	26.15	(5.70)	28.33	(4.24)	-1.22
	High	31.00	(4.74)	29.85	(4.16)	.66
8th	Low	25.43	(4.60)	27.40	(4.22)	-1.19
	High	29.60	(4.22)	30.35	(3.43)	-.46
2. The Sort Task						
6th	Low	8.92	(2.25)	8.25	(2.77)	.52
	High	10.43	(4.27)	9.50	(2.54)	.73
7th	Low	7.46	(3.15)	9.36	(3.98)	-1.50
	High	10.08	(2.61)	10.77	(3.77)	-.57
8th	Low	8.71	(3.43)	9.20	(3.63)	-.41
	High	10.40	(3.42)	10.00	(2.89)	.33
3. The Open-ended Guidance Question in Writing						
6th	Low	2.84	(1.28)	2.23	(1.24)	1.04
	High	3.29	(2.09)	3.15	(1.28)	.23
7th	Low	3.08	(1.32)	3.38	(1.26)	-.52
	High	4.77	(1.54)	4.85	(1.68)	-.13
8th	Low	4.50	(1.99)	3.73	(1.16)	1.36
	High	5.47	(1.30)	4.57	(1.55)	1.59

1. Index of Writing Awareness: MSW = 19.82
2. Sort Task: MSW = 10.41
3. Guidance Question in Writing: MSW = 2.29

*$p < .05$; $\underline{t}D = 1.87$

Table 5

Means, Standard Deviations, and Test Statistics for the Planned Grade Comparisons Within the Index of Writing Awareness, the Sort Task, and the Open-Ended Questions in Writing and Reading Measures

1. The Index of Writing Awareness:

Grade	Mean	SD	Grade	Mean	SD	Test Stat
8th	28.20	(4.47)	6th	26.30	(4.58)	2.25*
8th	28.20	(4.47)	7th	28.84	(4.97)	-.75
7th	28.84	(4.97)	6th	26.30	(4.58)	2.91*

2. The Open-Ended Guidance Question in Writing:

Grade	Mean	SD	Grade	Mean	SD	Test Stat
8th	4.56	(1.61)	6th	2.89	(1.54)	5.84*
8th	4.56	(1.61)	7th	4.02	(1.63)	1.90
7th	4.02	(1.63)	6th	2.89	(1.54)	3.83*

3. The Open-Ended Question in Reading:

Grade	Mean	SD	Grade	Mean	SD	Test Stat
8th	1.48	(.80)	6th	1.17	(.78)	2.20*
8th	1.48	(.80)	7th	1.44	(.83)	.28
7th	1.44	(.83)	6th	1.17	(.78)	1.87

1. The Index of Writing Awareness: MSW = 19.82
   n: 6th = 51; 7th = 51; 8th = 58

2. Guidance in Writing Question: MSW = 2.29
   n: 6th=53; 7th=52; 8th=58

4. Reading Question: MSW = .551;
   n: 6th=53; 7th=52; 8th=58

*p < .05; tD = 2.13

Students in the treatment group were predicted to increase the number of repositions and meaningful revisions within their papers while writing paper three (the prompted condition). In addition, because the computerized prompts were expected to focus the treatment group primarily on generating and evaluating content, it was postulated that they would make less surface changes/100 words of content generated as compared to the control condition. Since a limited number of papers were keystroke analyzed (60 total papers), grade by group planned comparisons were performed to increase the power of the analyses, instead of the grade by group by ability comparisons that were performed on previous measures. Due to disk failures in the keystroke mapping files near the end of the study, only paper one (pretreatment) and paper three (mid-treatment) were available for keystroke analysis; hence difference scores were obtained for these revision data.

Difference scores between the pretest (paper one) and the prompted condition (paper three) were created for: (1) surface changes (e.g., grammar, format, spelling, and word changes); (2) meaningful changes (e.g., phrase, sentence, and multisentence changes); and (3) page repositions/scrolling within a paper. No differences were predicted for the fourth measure, errant keystrokes, as a result of the treatment. The difference scores for each measure were used to perform treatment versus control planned comparisons at each grade level. Though none of the original directional Dunn grade by group comparisons were significant, the mean surface revision scores for the treatment

group within each grade were favorable (lower) when compared to their respective controls. Hence, the treatment subjects seem to have reduced the number of surface changes they made when writing in the prompted condition as compared to both their pretest scores and the control group, although, again, this was not statistically significant. Moreover, the treatment subjects appeared to have increased their amount of textual scrolling or repositions as compared to both their pretest scores and the control group.

In order to increase the power of the calculations and investigate possible global group differences (instead of grade/group), for each of these measures the appropriate 2 (groups) by 3 (grades) ANOVAs were executed. The ANOVA for surface changes confirmed that there was a main effect for group ($F_{1,48} = 5.88$, $p < .05$; see Table 6). The treatment group did make less surface levels changes per 100 words of text produced when exposed to the prompts than their pretest scores [Treatment: X = -2.89 (less changes in paper three), SD = 4.05; Control: X = -.35, SD = 3.96]. The ANOVA for reposition changes failed to find a main effect for group ($F_{1,48} = 3.11$, $p > .05$). However, because this test approached significance ($p < .10$), an additional post hoc ANOVA was performed on the reposition scores within paper three only (see Table 7). Here, a main effect for group was found ($F_{1,48} = 7.77$, $p < .05$). The treatment group was repositioning within their texts more than the control group when they were in the prompted condition (Treatment: X = 5.76 repositions, SD = 8.70; Control: X = 3.77, SD = 5.62). In addition, a two-way interaction between group and

ability was detected by this ANOVA ($F_{1,48} = 5.67$). The significant interaction occurred because high-ability treatment subjects appeared to respond to the prompts by scrolling through their texts more than low-ability treatment students (High-ability: X = 15.4, SD = 12.73; Low-ability: X = 5.47; SD = 6.85). In contrast, low-ability control students scrolled through their texts slightly more often than high-ability students, though for neither of these control groups did scrolling exceed five times per paper.

Older and more able students were expected to make fewer errant keystrokes and surface changes, while, in turn, generating more meaning changes and repositions within their texts. Table 8 consolidates the planned grade comparisons for three of the four primary measures obtained from the keystroke mapping and analyses: (1) errant keystrokes (a measure of student familiarity with the word processor and also the amount of time spent revising at the word level); (2) surface changes; and (3) paper repositions. As anticipated, sixth-graders concentrated their revisionary strategies more on correcting/altering errant keystrokes and the surface features of their texts per 100 words produced as compared to both eighth-graders (errant keys: $\underline{t}D = 3.32$, $p < .05$; surface changes: $\underline{t}D = 4.16$, $p < .05$), and seventh-graders (errant keys: $\underline{t}D = 2.92$, $p < .05$; surface changes: $\underline{t}D = 2.61$, $p < .05$). Though eighth-graders exhibited fewer keystroke and surface level revisions than seventh-graders per 100 words of text produced, pairwise comparisons between these two grades were not significant ($\underline{t}D = .40$, $p > .05$; and $\underline{t}D = 1.53$, $p > .05$, respectively). Though not shown in this

table, there were no significant differences between the three grade levels in meaningful revisions and repositions on paper three. Also, lower-ability students made significantly more errant keystrokes, surface changes, and less cursor repositions, but did not produce fewer meaningful changes per 100 words of text (see Table 9).

Two secondary measures obtained from the keystroke analyses were percentages of: (1) prompt effectiveness (prompts causing meaningful changes in text); and (2) apparent prompt ineffectiveness (prompts selected in succession without any textual change or cursor movement; multi-prompting). Though the prompts assisted sixth-graders nearly 20 percent of the time in creating meaningful changes, as compared to approximately 11 and 12 percent for the seventh and eighth-graders, respectively, the pairwise comparisons were not significant (6th vs. 7th: $\underline{t}TK = 2.08$, $p > .05$; 6th vs. 8th: $\underline{t}TK = 1.77$, $p > .05$). Similarly, prompts invoked in repetition without anything transpiring within the text (multi-prompting), occurred 47% of the time for the eighth-graders, 54.6% for the seventh-graders, and 36.5% for the sixth-graders. Post hoc Tukey comparisons were not significant on this measure. It must be pointed out for future research that planned directional comparisons favoring sixth-graders would have produced significant results when compared to seventh-grade on both prompt effectiveness and ineffectiveness. Though the results of earlier hypotheses would have provided insight into the direction of the differences in prompt effectiveness, such differences were not originally predicted

Table 6

Analysis of Variance Summary Table for Surface Level Difference Scores Between Papers One and Three

Source of Variation	Sum of Squares	DF	Mean Square	F
Grade	102.22	2	51.06	3.11
Group	96.55	1	96.55	5.88*
Ability	8.84	1	8.84	.53
Grade X Group	8.50	2	4.25	.26
Grade X Ability	4.23	2	2.14	.13
Group X Ability	14.07	1	14.07	.86
Grade X Grp X Abil.	5.88	2	2.94	.18
Residual	787.60	48	16.41	
Total	1027.82	59	17.42	

*p < .05; $F_{1,48}(.05) = 4.05$
**p < .05; $F_{2,48}(.05) = 3.20$

Table 7

Analysis of Variance Summary Table for Reposition Scores in Paper Three

Source of Variation	Sum of Squares	DF	Mean Square	F
Grade	372.03	2	186.02	2.34
Group	614.40	1	614.40	7.77*
Ability	299.27	1	199.27	3.78
Grade X Group	72.10	2	36.05	.47
Grade X Ability	249.03	2	124.52	1.57
Group X Ability	448.27	1	448.27	5.67*
Grd X Grp X Abil.	15.63	2	7.82	.10
Residual	3798.00	48	79.13	
Total	5868.73	59	99.47	

*p < .05; $F_{1,48}(.05) = 4.05$
**p < .05; $F_{2,48}(.05) = 3.20$

Table 8

Means, Standard Deviations, and Test Statistics for the Planned Grade Comparisons Within the Errant Keystroke, Surface Change, Meaning Change, and Reposition Measures

1. Errant Keystrokes: (younger > mistakes than older)

Grade	Mean	SD	Grade	Mean	SD	Test Stat
6th	19.95	(9.64)	8th	12.07	(4.98)	3.32*
6th	19.95	(9.64)	7th	13.03	(7.59)	2.92*
7th	13.03	(7.59)	8th	12.07	(4.98)	.40

2. Surface Changes: (younger > changes than older)

Grade	Mean	SD	Grade	Mean	SD	Test Stat
6th	6.26	(5.20)	8th	4.55	(3.44)	4.16*
6th	6.26	(5.20)	7th	5.19	(3.03)	2.61*
7th	5.19	(3.03)	8th	4.55	(3.44)	1.53

4. Repositions: (older > movement than younger)

Grade	Mean	SD	Grade	Mean	SD	Test Stat
8th	10.55	(10.77)	6th	4.55	(8.27)	2.13
8th	10.55	(10.77)	7th	6.60	(10.24)	1.40
7th	6.60	(10.24)	6th	4.55	(8.27)	.73

1. Errant Keystrokes: MSW = 56.33
2. Surface Changes: MSW = 16.87
3. Repositions: MSW = 79.13

n: 6th = 20; 7th = 20; 8th = 20

*p < .05; tD = 2.16 (Note: means are per 100 words of text generated)

Table 9

Means, Standard Deviations, and Test Statistics for the Post Hoc Ability Comparisons Within the Errant Keystroke, Surface Change, Meaning Change, Repositions, Prompt Effectiveness, and Prompt Ineffectiveness Measures

Measure	Low Ability Mean	SD	High Ability Mean	SD	Test Stat
Errant	17.45	(8.87)	12.61	(7.06)	2.49*
Surface	6.42	(4.99)	4.25	(2.28)	2.05*
Meaning	1.87	(1.95)	1.69	(1.56)	-.41
Reposition	9.47	(11.56)	5.00	(7.64)	1.95*
Effect	.14	(.08)	.16	(.11)	-.17
No Effect	.55	(.19)	.35	(.22)	2.66**

1. Errant Keystrokes: MSW = 56.33
   n: 6th = 20; 7th = 20; 8th = 20

2. Surface Changes: MSW = 16.87
   n: 6th = 20; 7th = 20; 8th = 20

3. Meaning Changes: MSW = 2.88
   n: 6th = 20; 7th = 20; 8th = 20

4. Repositions: MSW = 79.15
   n: 6th = 20; 7th = 20; 8th = 20

5. % of Prompts Effective: MSW = .009
   n: 6th = 10; 7th = 10; 8th = 8

6. % of Prompts Ineffective: MSW = .039
   n: 6th = 10; 7th = 10; 8th = 8

*p < .05; t = 1.68
**p < .05; tTK = 2.53

(see Bonk (1989) for further data on prompt effectiveness and ineffectiveness by grade and ability levels).

A number of additional correlational analyses and unplanned analyses were performed on this data. Meaningful revision scores were not correlated with the IWA, the holistic scores, or repositions within text. However, as predicted, surface changes in paper three were inversely correlated with students' holistic scores ($r = -.45$, $p < .001$). Next, scrolling/repositioning within text was significantly correlated with the holistic scores ($r = .49$, $p < .001$). Finally, the holistic scores for paper three and the IWA were significantly correlated ($r = .33$, $p < .001$).

In addressing possible attitude differences between groups toward using the computer as a writing tool, post hoc ANOVAs were performed on the 10 items of the computer questionnaire. ANOVA's for items 3 and 7 displayed main effects for group; $F1,148 = 4.07$, $p < .05$; and $F1,150 = 4.40$, $p < .05$). The treatment group responded more positively to question three (regarding whether the computer helped them to think about the reader more than when they wrote with pen and paper) (Treatment: $X = 4.78$; $sd = 2.83$; Control: $X = 3.84$, $sd = 2.61$). Similar results were obtained for question seven (regarding whether there were things they could write, think about, or do with the computer that they could not do with pen and paper) (Treatment: $X = 4.54$; $sd = 3.02$; Control: $X = 3.51$, $sd = 3.02$). Main effects for group were not significant on the other eight items of the questionnaire.

Summary of the Results/Implications

Writing is a complex problem solving process. It is even more complex for young children who lack a monitoring system that enables them to switch effectively between the generative and evaluative demands of both planning and revising text. The generative and evaluative intervention used in this study was one attempt to influence children's metacognitive control of their writing plans and goals, thereby attempting to overcome their usually poor content generation coping strategies. The main source of support for the treatment program devised was in the revisionary strategies undertaken by students exposed to the prompts. However, evidence supporting the usefulness of thinking skill prompts within either a particular paper or over a time (in this case six weeks) was minimal. Perhaps students' content generation coping strategies were not easy habits to change. Still, in a short-term intervention with college students using this generative/evaluative prompting framework, Reynolds (1990) found treatment-related increases in both process (meaningful revisionary practices) and product (overall writing quality) variables.

In this study, development was not only examined between groups and over time, but also was explored between three grade levels. In most measures there was a significant main effect for grade that favored the older students. Within the holistic score analyses of papers three and five, the only comparison lacking in significance was between the seventh- and eighth-graders on paper five (the posttest). The grade comparisons performed on the IWA, the open-

ended question in writing, errant keystrokes, and surface change scores provided consistent findings: both eighth-graders and seventh-graders performed better than sixth-graders, but there were no differences between the seventh- and eighth-graders.

What might be concluded from these coinciding findings? One possible conclusion is that there was a significant increase in metacognitive skill within writing that originated somewhere between sixth- and seventh-grade, and manifested itself in improved writing performance. Such a conclusion may account for why the sixth-graders had the most positive attitudes toward the project, and why they made most effective (though not significant) use of the prompts; they lacked the internal metacognitive cuing system which the computer provided for them. Alternatively, one could argue that since one person teaches both seventh- and eighth-grade English, the older students were exposed to different writing skills and strategies than the sixth-graders. Thus, the sixth- to seventh-grade shift in writing skills and strategies must be replicated, especially since much of the previous research has suggested that writing skills improve around the eighth grade, not sixth or seventh (Bereiter and Scardamalia, 1987). Additional analyses of the IWA will address this issue more directly (Bonk, in preparation).

The lack of positive effects of the prompts on students' writing performance should not immediately discount their utility. The fact that student performance generally did not decrease in this novel writing environment suggests that students can readily adapt to different types of word processing environments. For example,

the layout of the prompts within the bottom two lines of the screen did not appear to inhibit performance. Children did not complain about the loss of two lines of text. In fact, the computer attitudes questionnaire indicated that the prompts helped the treatment group think of potential readers and accomplish more things than they could with pen and paper.

Prompt effectiveness was investigated by grade and ability level. High-ability children made some type of change or textual movement after 65% of the prompts, while low-ability students made changes only 45% of the time. Prompt effectiveness or productiveness (defined as prompts causing meaningful changes within text) ranged from about 10-12% for the seventh and eighth-graders to almost 20% for the sixth-graders. Though these differences were not significant, they, nevertheless, favored the younger students. In addition, while the percentages of prompt effectiveness were not astoundingly high, something noteworthy occurred within most students' texts as a result of the prompts--a meaningful addition or deletion--that may not have occurred without the prompts. However, if a student used 15-20 prompts during the course of two weeks (the time spent writing one paper), a 15% effective rate would result in only 2.6 meaningful textual changes. How much would performance on a 200-400 word paper improve from 2-3 additional phrases or sentences? In this particular study, the improvement was negligible.

Through the testing a more inclusive model of procedural facilitation in writing was developed and tested, a

generative/evaluative framework for prompted writing, changes in students' revisionary patterns as a result of access to the prompts was discovered. Students in the prompted condition were producing fewer surface level changes when exposed to the prompts, as compared to both their earlier performance and the control group. In addition, the treatment group increased the number of repositions they made within text when prompts were available in paper three. Apparently, the prompts caused the treatment subjects to scan their text more frequently than the control subjects. As indicated by a two-way interaction, this was especially true of high-ability subjects in the treatment condition. Nevertheless, the combination of reduced time spent on surface level aspects of text and the additional time allocated to the scrolling of one's text did not result in either more meaningful changes or better quality texts. Maybe the students were pondering over the prompts, but did not know how to respond to them. However, if that were true, one would have predicted that the older students would have made more effective use of the prompts than the younger students. A more plausible explanation is that the categorization scheme used to code meaningful changes is in need of restructuring (see Reynolds and Bonk, 1990).

Although the treatment group performed below original expectations, progress was made in understanding the connection between writing and thinking, as reflected in correlations among the instruments used to measure metacognition, internalization, and writing performance. Two of the strongest correlations were found

between the IWA and the holistic scores, and between the open-ended question in writing and the holistic scores. Thus, it appears that an understanding of one's writing strategies has an impact on the quality of text produced. Students who have an understanding of how to plan, evaluate, regulate, and conditionally apply writing strategies appear to produce higher quality texts. Though this is intuitively simple, the impact that the IWA may have for diagnosticians attempting to measure the success of their metacognitive interventions, and for English teachers searching for an easily administered diagnostic tool, could be significant. As noted earlier, additional work is warranted (Bonk, in preparation).

There are a number of treatment variations from the current program that might be utilized to further explore the metacognitive and performance effects of prompts within children's writing. For instance, students in future treatment groups could receive greater initial modeling of the prompt categories. In addition, student involvement in the creation of prompts and prompt categories along with the use of the prompts for a semester or academic year also could eliminate some of the problems encountered here.[4] Yet, before commencing in either of these directions, researchers must ask themselves if the extra time spent within an individual classroom or school would be fruitful. Is their enough merit in the ideas behind procedural facilitation and internalization of prompts to

---

[4]Scott Paris (personal communication) concluded that widespread implementation of pragmatic treatments like this one would require one year or more for significant variation to be detectable across classes.

warrant the continuation of this research?

With the keystroke mapping revision data that was lost to disk failure (paper number five), time-consuming keystroke mapping and revision scoring must also have a high payback in order to be considered worthwhile. Nevertheless, keystroke analyses enabled the experimenter to document when and where prompts were most effective. For instance, sixth-graders and high-ability students were more productive in using the prompts than older and lower-ability students, respectively. Additional investigations should be made regarding why prompts may be more useful at certain grades, ability levels, and moments in the composing process. Several other studies might be undertaken to increase our understanding of the effectiveness of prompts within writing. One possible study might focus on the usefulness of prompts at the sixth-grade level using multiple treatment groups (e.g., one group receiving generic generative and evaluative prompts used in the present study, a second group receiving topic-specific prompts, and a third treatment group creating and using their own prompt categories and prompts).

The present study also illustrates that the computer tools (and other media) cannot be separated from instruction. Effective writing instruction suggests that introducing the objectives and principles of the prompts in the classroom prior to encountering them on the computer might facilitate their usage (Hillocks, 1986). Studies that integrate new prompting media with appropriate instruction should increase the possibility of their internalization. Moreover, studies that increase our understanding

of how to facilitate thought processes involved in superior writing performance are critical to the evolution of writing research. At the present time, however, few studies meet that goal. The current study was one attempt to expand the envelope of what is known about the connection between writing and reasoning. The additional studies proposed above (e.g., those involving greater integration with classroom instruction, modeling, student involvement, and peer feedback) should further this understanding.

References

Applebee, A. N. (1982). Writing and learning in school settings. In M. Nystrand (Ed.), What writer's know: The language, process, and structure of written discourse. New York, NY: Academic Press, Inc.

Applebee, A. N. (1984). Writing and reasoning. Review of Educational Research, 54(4), 577-596.

Barritt, L. S., & Kroll, B. M. (1978). Some implications of cognitive-developmental psychology for research in composing. In C. R. Cooper, & C. R. Odell (Eds.), Research on composing. Urbana, IL: National Council of Teachers of English.

Bartlett, E. J. (1982). Learning to revise: Some component processes. In M. Nystrand (Ed.), What writer's know: The language, process, and structure of written discourse. New York, NY: Academic Press, Inc.

Bereiter, C. (1980). Development in writing. In L. W. Gregg, & E. R. Steinberg (Eds.), Cognitive processes in writing. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bereiter, C., & Scardamalia, M. (1985). Cognitive coping strategies and the problem of "inert knowledge." In J. W. Segal, S. F. Chipman, & R. Glaser (Eds.), Thinking and Learning Skills (Vol. 1, Research and Open Questions). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bereiter, C., & Scardamalia, M. (1987). The psychology of written communication. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bonk, C. J. (1989). The effects of generative and evaluative computerized prompting strategies on the development of children's writing awareness and performance. Dissertation Abstracts International, (University Microfilms No. 8923313).

Bonk, C. J. (1990). A synthesis of social-cognition and writing research. Written Communication, 7(1), 136-163.

Bonk, C. J. (in preparation). The Index of Writing Awareness: One measure of metacognitive ability in writing. Unpublished manuscript, West Virginia University.

Bridwell, L. S., Nancarrow, P. R., & Ross, D. (1984). The writing process and the writing machine: Current research on word processors relevant to the teaching of composition. In A. Beach, & L. S. Bridwell (Eds.), New directions in composition research. New York, NY: The Guilford Press.

Bridwell, L., Sirc, G., & Brooke, R. (1985). Revising and computing:

Case studies of student writers. In S. W. Freedman (Ed.), The acquisition of written language. Norwood, NJ: Ablex Publishing Company.

Bruce, B., Collins, A., Rubin, A. D., & Gentner, D. (1982). Three perspectives on writing. Educational Psychologist, 17(3), 131-145.

Caccamise, D. J. (1987). Idea generation in writing. In Matsuhashi, A. (Ed.), Writing in real time: Modelling production processes, Norwood, NJ: Ablex Publishing Corporation.

Carey, L. (1988). Pushing planning to the limit: Using computer based prompts to elicit writers' plans. Paper presented at the AERA annual convention, New Orleans, LA.

Collins, A., & Gentner, D. (1980). A framework for a cognitive theory of writing. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive processes in writing. Hillsdale, NY: Lawrence Erlbaum Associates, Inc.

Comeaux, M., & Lehrer, R. (1986). The mind of a writer: A cognitive model of computer-assisted composition. Paper presented at the AERA annual convention, San Francisco, CA.

Daiute, C. (1985). Do writers talk to themselves? In S. W. Freedman (Ed.), The acquisition of written language: Response and revision. Norwood, NJ: Ablex Publishing Corporation.

Daiute, C. A. (1986). Physical and cognitive factors in revising: Insights from studies in computers. Research in the Teaching of English, 20(2), 141-159.

Davis, G. A. (1986). Creativity is forever. (2nd ed.). Dubuque, IA: Kendall/Hunt Publishing Company

DiPardo, A., & Freedman, S. W. (1988). Peer response groups in the writing classroom: Theoretical foundations and new directions. Review of Educational Research, 58(2), 119-149.

Englert, C. S., & Raphael, T. E. (1988). Constructing well-formed prose: Process, structure, and metacognitive knowledge. Exceptional Children, 54(6), 513-520.

Englert, C. S., Raphael, T. E., Fear, K. L., & Anderson, L. M. (1988). Students' metacognitive knowledge about how to write information texts. Learning Disability Quarterly, 11, 18-46.

Faigley, L., & Witte, S. (1981). Analyzing revision. College Composition and Communication, 33(4), 400-414.

Fitzgerald, J. (1987). Research on revision in writing. Review of Educational Research, 57(4), 481-506.

Fitzgerald, J., & Markham, L. R. (1987). Teaching children about revision in writing. Cognition and Instruction, 4(1), 3-24.

Flower, L. (1987). The role of task representation in reading to write. Berkeley, CA & Pittsburgh, PA: Center for the Study of Writing, University of California & Carnegie Mellon University, Technical report No. 6.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communication, 32, 365-387.

Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, diagnosis, and the strategies of revision. College Composition and Communication, 37(1), 16-55.

Freedman, S. W. (1987). Response to student writing. Urbana, IL: National Council of Teachers of English.

Freedman, S. W., Dyson, A. H., Flower, L., & Chafe, W. (1987). Research in writing: Past, present, and future. Berkeley, CA & Pittsburg, PA: Center for the Study of Writing, University of California & Carnegie Mellon University, Technical Report No. 1.

Graham, S., & Harris, K. R. (1988). Instructional recommendations for teaching writing to exceptional students. Exceptional Children, 54(6), 506-512.

Haas, C. (1988). How word processors affects planning in writing: The impact of technology. Paper presented at the AERA annual convention, New Orleans, LA.

Hawisher, G. E., & Fortune, R. (1988). Research into word processing and the basic writer. Paper presented at the AERA annual convention, New Orleans, LA.

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. American Psychologist, 1106-1113.

Hillocks, G. (1986). Research on written composition: New directions for teaching. Urbana, IL: National Conference of Research in English & ERIC Clearinghouse on Reading and Communication Skills.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues and definition, measurement, and instruction. Educational Psychologist, 22(3 & 4), 255-278.

Martlew, M. (1983). Problems and difficulties: Cognitive and communicative aspects of writing development. In M. Martlew (Ed.), The psychology of written language: Developmental and educational perspectives. New York, NY: John Wiley & Sons.

McCutchen D. (1988). "Functional automaticity" in children's writing: A problem in metacognitive control. Written Communication, 5(3), 306-324.

Paris, S. G. (1988). Fusing skill and will in children's learning and schooling. Paper presented at the AERA annual convention, New Orleans, LA.

Reynolds, T. H. (1990). Facilitating college student's textual reprocessing through generate/evaluate iterations. Paper presented at the AERA annual convention, Boston, MA.

Reynolds, T. H., & Bonk, C. J. (1990). Windows on writing: The usefulness of keystroke mapping to monitor writing progress. Paper presented at the 6th Computers and Writing Conference, Austin, TX.

Ruth, L., & Murphy, S. (1988). Designing writing tasks for the assignment of writing. Norwood, NJ: Ablex Publishing Corporation.

Salomon, G. (1988). AI in reverse: Computer tools that turn cognitive. Journal of Educational Computing Research, 4(2), 123-139.

Scardamalia, M., & Bereiter, C. (1982). Assimilative processes in composition planning. Educational Psychologist, 17(3), 165-171.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In J. W. Segal, S. F. Chipman, & R. Glaser (Eds.), Thinking and learning skills (Vol. 1, Research and open questions). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In M. C. Wittrock (Ed.), Handbook of research on teaching. (3rd edition, pp. 778-803). New York: Macmillan Education Ltd.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. Cognitive Science, 8, 173-190.

Talbot, J. (1986). The assessment of critical thinking in history/social science through writing. Social Studies Review, 25(2), 33-41.

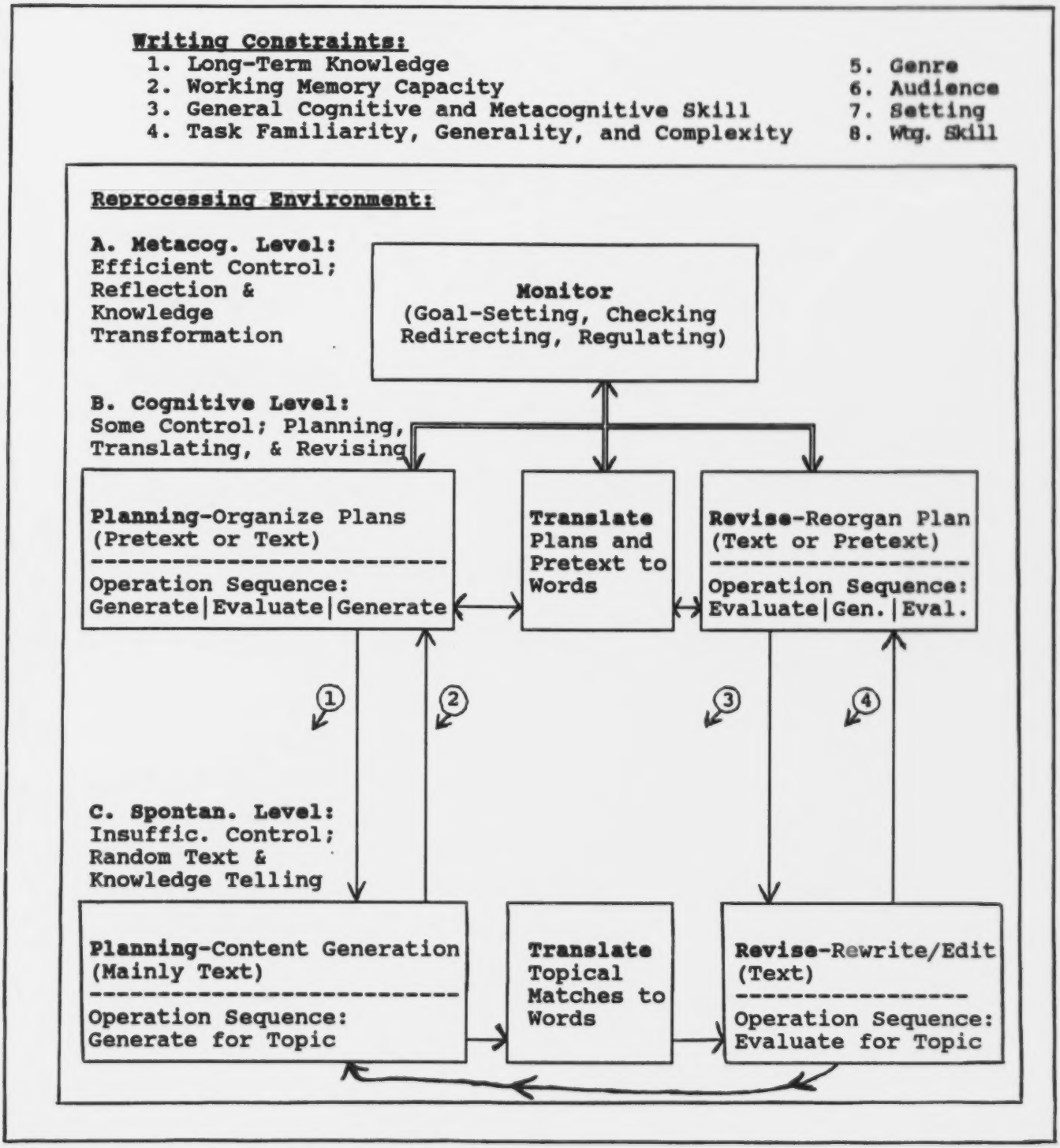
Torrance, E. P. (1974). Norms and technical manual: Torrance tests of creative thinking (rev. ed.). Bensenville, IL: Scholastic Testing Service.

Vygotsky, L. (1986). Thought and language (rev. ed.). Cambridge, MA: The MIT Press.

Woodruff, E., Bereiter, C., & Scardamalia, M. (1981). On the road to computer assisted composition. Journal of Educational Technology Systems, 10(2), 133-148.

Appendix A

Proposed Model of the Generative/Evaluative Processes in Composing


Writing Constraints:
1. Long-Term Knowledge
2. Working Memory Capacity
3. General Cognitive and Metacognitive Skill
4. Task Familiarity, Generality, and Complexity
5. Genre
6. Audience
7. Setting
8. Wtg. Skill
Reprocessing Environment:
A. Metacog. Level: Efficient Control; Reflection & Knowledge Transformation
Monitor (Goal-Setting, Checking Redirecting, Regulating)
B. Cognitive Level: Some Control; Planning, Translating, & Revising
Planning-Organize Plans (Pretext or Text)
Operation Sequence: Generate|Evaluate|Generate
Translate Plans and Pretext to Words
Revise-Reorgan Plan (Text or Pretext)
Operation Sequence: Evaluate|Gen.|Eval.
1
2
3
4
C. Spontan. Level: Insuffic. Control; Random Text & Knowledge Telling
Planning-Content Generation (Mainly Text)
Operation Sequence: Generate for Topic
Translate Topical Matches to Words
Revise-Rewrite/Edit (Text)
Operation Sequence: Evaluate for Topic

1. Downsliding from reflective processes toward automaticity or fluency.
2. Movement toward reflection after prompted generative processes.
3. Downsliding from reflective processes to edit and make mechanical changes.
4. Movement toward reflection after prompted evaluative processes.

Appendix B

Generative and Evaluative Prompt Listing

Generative Prompts

MORE IDEAS:

_____ alt/shift Q: List all that you know about your topic in your head or on paper. You may want to jot down items in the list that are not in your paper.

_____ alt/shift A: Ask yourself: What other ideas does this suggest? What could I add here? And, how could I exaggerate or maybe say the opposite?

_____ alt/shift Z: What else might your audience want to know? Would the reader want to know about the smell, sight, sound, or touch of your object?

TYPES OF IDEAS:

_____ alt/shift W: Add other categories, models, examples, or lists. You might try to use pictures in your head to compare points.

_____ alt/shift S: Just imagine if everything you've said so far is wrong. If the reader caught it, what changes might he/she suggest?

_____ alt/shift X: Think again about your reader. Are there other points of view that are necessary for your reader to understand?

NEW IDEAS:

_____ alt/shift E: Try out a wild idea or describe your last thought in a metaphor. How is a ____ like a ____???

_____ alt/shift D: Ask yourself "What if...?" and then reflect on what might happen to change your mind on this topic.

_____ alt/shift C: Try combining two or more of your ideas into something really unique. Have you used your creativity or imagination?

EXTENDERS:

_____ alt/shift R: Have some fun, play with the last idea, expand or extend it, and then maybe contrast it with something else.

_____ alt/shift F: Reread your last paragraph. Would expanding or adding a sentence help your reader understand?

_____ alt/shift V: Try to broaden the focus of your paper by including exceptions to what you are saying.

Evaluative Prompts

QUALITY:

_____ alt/shift T: Think about the problem or original topic. Is everything you've said needed or related to it?

_____ alt/shift G: Reread your paper and delete repeated or unneeded sentences which don't help form an overall theme.

_____ alt/shift B: Try to see or imagine where your writing is headed. Is the information you're providing good and also relevant?

CLEAR/LOGICAL:

_____ alt/shift Y: Give an example that might make your reasoning clearer to the reader. State all examples in clear and simple ways.

_____ alt/shift H: Read the first and last sentences to each paragraph. Are there transitions from one sentence to the next?

_____ alt/shift N: Think back about your original idea or opinion on this topic. What can you say now to provide support for your entire paper?

ASSUMING:

_____ alt/shift U: Reflect on the sources of your information. Are your sources and your assumptions stated as such in your paper?

_____ alt/shift J: Read over your paper for personal bias; look out for sentences where you say "I feel" or "I think" without backing them up.

_____ alt/shift M: Will your audience agree with your values, opinions, or ideas? If not, list something that might help get your point across.

CONCLUSIONS:

_____ alt/shift I: Have you provided enough information to back up your claims and conclusions? And are there other effects to what you're saying?

_____ alt/shift K: Are there different conclusions to what you are saying? Try to explain these so they make sense for the reader.

_____ alt/shift O: Can you summarize to the reader what you have said in one or two sentences? Try to do this at the end of each paragraph or idea.

OTHER PROMPTS (for whole paper):

_____ alt/shift L: Step back and look at your whole paper. Are your thoughts and ideas logically stated, justified, interesting, and unique?

_____ alt/shift P: (Clears the bottom screen/prompt window)

Appendix C

Post-Treatment Writing and Reading Metacognitive Questionnaire

---

Part A. Writing: Assume that you have a close friend who does not "write" very well...

1. What advice would you give him or her to help overcome this writing problem (Salomon, 1988)?

Scoring Points: (sample responses)

0. no strategy; no answer.
1. resemblance of a strategy (poorly worded answer); focus on handwriting.
2. more than one strategy listed, but not the best.
3. linking/connecting ideas; major rewrite; questioning strats.

2. What steps should he/she follow in writing his/her next report?

0. no answer; no strategy.
1. generate ideas and notes.
2. generate and organize notes.
3. generates and evaluate text written; provide support for main ideas and offer concluding sentences.

3. What would you tell him/her is a sign of good writing?

0. no answer; no strategy.
1. resemblance of a strategy (poorly worded);
2. an advanced/more formed strategy; (e.g., imagining where one's writing is headed)
3. focus on: audience awareness; clear logical flow; good story ending; transitions; other; (lists > than one strategy)

Part B. Reading: Assume that you have a close friend who does not "read" very well...

1. What advice would you give him or her to help overcome this reading problem (Salomon, 1988)?

Scoring: (sample responses)

0. no strategy; no answer.
1. strategy listed is of limited utility, or based on classroom knowledge (e.g., practice).
2. more than one common knowledge strategy listed (e.g., practice plus getting help).
3. a more advanced strategy is listed (e.g., getting a regular tutor, looking at the context, change one's environment, ignoring the details and trying to make global inferences; flipping through the text and searching for summary questions and headings).
4. identifies comprehension as the possible problem and suggests a strategy to help.
5. lists a couple of strategies or mentions questioning oneself.

Appendix D

Post-Treatment Computer Attitudes Questionnaire

Students will rate the following questions on a scale from 0 to 9, with `0' being low or definitely not and `9' being high or definitely yes. A `4' would be slightly negative and a `5' slightly positive.

1. Did using the computer **help you plan or organize** your ideas or notes any better than pencil and paper?

2. Did using the computer help you **come up with less ideas** than you normally have with pencil and paper?

3. Did the computer help you to **think about the reader more** than when you write with pen and paper?

4. Did the computer **help you to revise more** than with pen and paper?

5. **How much were you thinking** when you were writing at the computer?

6. Overall, did the computer system **make it harder** to write?

7. **Were there things** you could write or think about with the computer or that you were able to **do with the computer that you could not do with pen and paper?**

8. Was the computer system **difficult** to use?

9. **Were you writing better before** this writing project started?

10. Would you like to a writing project like this **next year**?

Appendix E

Index of Writing Awareness

This instrument was used to measure metacognition in writing. The format of the instument was based on the Index of Reading Awareness from Jacobs & Paris (1987).

Score:

Evaluation Component

1. What is the hardest part of writing for you?
1 a. Using complex words and sentences.
2 b. Organizing the paper in a general theme.
0 c. There is never anything difficult about writing for me.

2. What would help you become a better writer?
1 a. If more people would help you when you write.
0 b. Writing papers with shorter words or phrases.
2 c. Making sure that others will understand what you write.

3. What might help the average person with his/her writing?
0 a. Having notepaper and pencils always available.
1 b. Having access to questions that guide his/her writing.
2 c. To know when to organize and when to extend his/her ideas.

4. What is special about the first sentence or two in every paper you write?
1 a. They use key words to inform the reader and capture his/her attention.
0 b. The first few sentences are the most interesting.
2 c. They often tell what the story or paragraph will be about.

5. How are the last sentences of a paragraph or paper you write special?
0 a. They are exciting and cover new topic areas.
2 b. They summarize thoughts and ideas for the reader.
1 c. They are harder to write and are sometimes neglected by the writer.

Planning Component

6. If your teacher gave you a hard topic to write about, what should you do?
0 a. Write about an easier topic.
1 b. Try to tell at least some of the main or basic ideas.
2 c. Ask yourself questions to get at what you know and do not know.

7. When you get a writing assignment, what is one of your primary concerns?
2 a. What is going to happen or be presented in the your story or paper.
0 b. How long the assignment must be.
1 c. Whether you can think of any interesting or exciting sentences for this paper.

8. If the teacher told you to write a story that would be easy for him/her to remember the general meaning, what would you do?
2 a. Have introductory and summary sentences in each paragraph.
1 b. Make it easy for him/her by writing only one or two summary sentences.
0 c. List as many details about the topic as you could possibly know.

9. Before you start to write, what kind of plans should you make to help you to write better?
0 a. A writer doesn't need plans, he/she can usually just start to write.
1 b. You could choose a comfortable place for coming up with ideas and then begin.
2 c. You could think about why you are writing about this and then ask yourself questions about the topic.

10. If you were writing well, but were thinking ahead that you might not have enough information for one or two key issues, what might you do?
2 a. Try to generate or create more information through internal questions.
0 b. Ignore that part of the assignment.
1 c. Try to cover yourself by writing more on other parts of the assignment.

Regulation Component

11. Why do you go back and rewrite or recopy things over again?
0 a. Because it is good practice.
2 b. Because other possible readers may not understand it.
1 c. Because you want to check your spelling, grammar, and sentence order.

12. Why would you ask yourself about the logic or quality of an idea before adding it into your paper?
2 a. Because every idea must be fit into the paper's overall theme and flow.
0 b. No reason; good ideas can be added anywhere in one's paper.
1 c. Because ideas must have a purpose.

13. Why does what you want to say in a paper or story ever change?
2 a. Because at any time you can think of new or related ideas.
1 b. Because you come up with more to say.
0 c. Because you didn't write enough for the teacher during the first draft.

14. What do you do if you don't know what the writing assignment means or what to write about?
1 a. Read the assignment over again for ideas.
0 b. Write what you can about anything that might be related.
2 c. Think about the purpose of the assignment or ask for help.

15. Why is writing different from story telling?
1 a. In writing, the sentences and paragraphs have to end somewhere.
2 b. There are certain assumptions you can make when talking to someone, but in writing you may have to answer more questions and explain more.
0 c. There is no difference, you say the same things.

Conditional Knowledge Component

16. If you were told that your neighbor who works at the Milwaukee Sentinel might publish your article on "how to improve any school's cafeteria plan" if it was good enough, what would you do?
1 a. Have a friend and possibly your teacher read it over before you send it in.
0 b. Like any paper, you would send out your ideas written in the order in which you thought of them.
2 c. Be creative in coming up with ideas, but also be critical of whether they really fit into what the Sentinel's readers might want to hear about.

17. If you are writing a paper for a science or social studies class, what should you do in order to cover the important information and write well?
2 a. Constantly ask yourself questions about the goal of the project and the reader's needs.
0 b. Skip writing about the words or phrases that you don't understand.
1 c. Concentrate on important issues and try hard to do your best.

18. If you are writing an article for a children's magazine, which would help the most?
1 a. Rewriting and reworking the article as many times as possible.
2 b. Talking about it with someone to make sure that they also understood it.
0 c. Making sure everything is spelled correctly.

19. If you are writing a paper for a class project, what is a valuable first step?
1 a. Go to the library for information.
2 b. List or write down your initial ideas and brief notes.
0 c. Skip the parts of the assignment that you don't understand.

20. Which of these is the best way to revise and change a paper or story?
0 a. Concentrate on finding the "right word(s)" and spelling everything correctly.
1 b. Think about suggested changes your teachers or friends might make.
2 c. Think about how paragraphs and sentences might be reordered and expanded.

Appendix F

Holistic Rating Scheme

Readers assigned scores based on a holistic scoring guide. Holistic scoring ranged from a low of '1' to a high of '6' according to the following categories:

Score	Guidelines
6	A "6" essay demonstrates a high degree of competence in writing, though it may have minor errors. A paper in this category - is well organized and well developed - uses appropriate details to support a thesis or illustrate ideas - shows unity, coherence, and progression - demonstrates syntactic variety - displays clear facility in the use of language
5	A "5" essay clearly demonstrates competence in writing, though it may have minor errors. A paper in this category - is well organized and well developed, though it may have fewer details than does a 6 paper - shows unity, coherence, and progression - demonstrates some syntactic variety - displays facility in language, though it may not be as fluent as a 6 paper
4	A "4" essay demonstrates competence in writing, though it may have occasional errors. A paper in this category - is adequately organized and developed - uses some details to support a thesis or illustrate ideas - demonstrates adequate facility with language - may contain occasional writing errors, but they will be neither serious nor frequent
3	A "3" essay may demonstrate some competence in writing, but it is clearly flawed. A paper in this category is depreciated by - inadequate organization or development - failure to support a thesis or illustrate generalizations with appropriate detail - lack of variety in sentence structure - limited or inappropriate word choice - a pattern or accumulation of errors in mechanics, usage, or sentence structure
2	A "2" essay suggests incompetence in writing. A paper in this category is marked by one or more of the following weaknesses - disorganization or very little development - little or no detail or irrelevant specifics - serious errors in mechanics, usage, or sentence structure
1	A "1" essay clearly demonstrates incompetence in writing. A paper in this category will contain - serious and persistent writing errors - may also be illogical, incoherent, or severely underdeveloped.

Source--Copyright 1984 by Educational Testing Service. Reprinted by permission of the author.

Appendix G

Adaptation of Purves' Dimensional Scoring Instrument

Dimensional Scoring

Scoring Guide: Rank each composition on each aspect below on a scale from 1 (low in this quality) to 5 (displays this quality in a highly appropriate fashion). A '3' would indicate adequate performance. Rank each aspect separately.

Generative Criteria:

1. Adequacy of information presented -- enough to fulfill assignment. Refers to the degree to which all of the relevant information from the stimulus is contained in the text. (Fluency--more ideas.)

2. Consideration of alternatives -- other arguments or interpretations and opinions of others. Refers to the extent to which the writer appears to admit the possibility of alternative or counter arguments or interpretations and accepts them as admissable or rebuts them. Includes implied rebuttles, qualifying language, hedging, modesty, and consideration of other viewpoints. (Flexibility--thinking of different approaches or categories.)

3. Inferences made beyond the scope of the assignment. Refers to the number and depth of interpretations (casual, resultant, comparative, contrastive, extrapolative) that the writer makes beyond the information in the stimulus and or from the outside. (Originality--uniqueness or nonconformity in thought; see also Elaboration.)

4. Richness of additional information -- relevant allusion (an implied or indirect reference). Refers to the use of additional information/details to that which is in the stimulus (e.g., information drawn from a variety of sources such as reading or general knowledge) and may be seen as the amount of relevant allusion. (Elaboration--adding details and embellishments.)

Evaluative Criteria:

1. Relationships drawn among items of information. Refers to the degree to which the text shows that connections have been made between the various items of information and the validity and/or complexity of the relationships drawn. (Relevancy--needed information that enhances overall theme; see also Logic.)

2. Framing/Grouping/Unity. Refers to the degree to which the writer presents the content in such a manner that there is an apparent beginning, middle, and end to both paragraphs and the paper. It also refers to the degree to which the writer joins the various pieces of information, relationships, or inferences into paragraphs or some other means. Inadequate writing would be denoted by extraneous matter and thoughts, combining bits of information, and lists with no discernible pattern or system. (Logic--clear flow of the paper; see also Relevancy.)

3. Evaluation -- the making of judgments as to the merit of ideas. Refers to the degree to which the writer appears to make judgments as to the relative merit of particular relationships, inferences, or synthesis and the degree to which applicable criteria are used. Includes implied and explicit reasons why suggest something, why change something, why build something, and why qualified to do something. (Assumptions--recognizing bias in one's own thoughts; see also Relevancy.)

4. Synthesis -- drawing together ideas into a generalization. Refers to the degree to which the writer appears to draw together information, relationships, and inferences into a single or complex generalization. (Conclusions--drawing together appropriate summary statements.)

Adapted from Purves' Analytical Scheme for Critical Thinking Skills, see Hawisher and Fortune (1988). Permission granted by the author to modify and reproduce.